Individual identification of red deer (*Cervus elaphus*) based on DNA extractions from faeces.

by
Liv Guro Kvernstuen

Cand. Scient. thesis

University of Oslo
Department of Biology
Centre for Ecological and Evolutionary Synthesis (CEES)
2004

Forord

Først vil jeg takke alle veilederne mine, ja for selv om dere har blitt mange, så har dere alle vært uunnværlige.
Nils Christian, du er en levende kilde av energi, en man alltid får et ærlig svar fra og en som alltid er der når man trenger deg. Tusen takk for at du ikke har mistet troen på meg.
Kjetill, fra deg har jeg alltid kunnet regne med gode ideer og innspill, du har alltid svart på mine spørsmål og dine kommentarer til oppgaven hadde jeg ikke greid meg uten. Takk for at du lot meg som zoolog delta i et så fascinerende prosjekt som har gitt meg en brennende interesse for molekylær økologi.
Atle, du har hele tiden stilt opp for meg og din dør har alltid vært åpen. Uten deg hadde jeg knekt sammen for lenge siden.
Eli, du har lært meg alt jeg kan av laboratoriearbeid, du har svar på de dummeste spørsmål og du tørket mine tårer når fortvilelsen tok meg. Hadde aldri holdt ut all tiden på labben hadde det ikke vært for deg.
Øystein, vel ankom du sist, men du kom da det var som mørkest. Uten deg ville det aldri blitt noen resultater, du er den som alltid ser en mulighet og du er en uhelbredelig optimist. Er det noen man kan regne med så er det deg, og hos meg vil du alltid ha en glorie.

Tusen takk til Norsk Hjortesenter på Svanøy for all den hjelpen jeg fikk til prøvetagning. Takk til daglig leder Johan Trygve Solheim, for hjelpa den påsken jeg var der, og til Vebjørn Veiberg for at du åpnet ditt hjem for meg.

Takk til alle som har vært på DNA-lab, og spesielt takk til Emelita Rivera Nerli. Og tusen takk til Jon Nissen-Meyer for at du lot meg få bruke labben din, selv om jeg på ingen måte er en biokjemiker.

Til den ”gamle” gjengen på Ørnereiret, takk for at dere har hørt på mine frustrasjoner og gledeshyl. Og takk til de som uforpliktet har lest igjennom oppgava mi, spesielt takk til Anna, Geir og Kirsti.

Lisbeth, takk for at du alltid har vært der, til deg blir ingen blomsterbukett stor nok. Aina, ingen problemer er for små for deg, du har alltid stilt opp med et smil som smelter is. Kirsti, du har vært der i snørr og tårer, aggresjon og sinne, glede og jubel, du er den vennen alle drømmer om. Trude, du sa jeg kom til å forbli på bio, og du fikk rett, takk for du har tatt deg av lille meg. Og alle dere andre som jeg har plaget med sorger og gleder igjennom studietida, tusen takk for at dere har holdt ut med meg. Dere er venner jeg aldri glemmer.

Kjære Geir, du kom inn i min verden sammen med biologien, og det var det beste som kunne hendt. Du har holdt ut med meg, stått ved min side i sorg og glede, og vært grunnpilaren i mitt liv siden vi fant hverandre. Takk for at du er deg og at du er glad i meg. Dette hadde jeg aldri klart uten deg.

Guro

Abstract

Recent advances in molecular genetic techniques have made it possible to genotype DNA extracts from non-invasive samples such as hair, faeces, and urine. Subsequent identification of individuals from such material has allowed estimation of important parameters such as population size, immigration rate, and reproductive contribution from immigrants, even without handling the animals. This thesis presents a pilot study assessing the reliability of using faecal samples for genotyping in red deer. The test was preformed on samples of blood and faeces from 42 farmed animals, using six microsatellites. The reliability of the method was assessed through a multiple-tube approach and by comparing faecal genotypes with the corresponding genotypes obtained from blood. DNA of sufficient quality to be genotyped for all six loci was extracted from 23 faeces samples (54.8 %) by a silica-based method. Allelic drop-out was found in 7.7 % of the amplifications from extracts of faeces. Comparison of blood and feces samples showed that correct consensus genotypes were obtained at all loci and samples already after the three first replicates. Probability of identity were estimated to PI = 6.5×10^{-6} for unrelated individuals and PI = 8.0×10^{-3} for siblings. These results suggest that DNA extracted from red deer faeces is a viable source for obtaining reliable individual genotypes, and that it can be achieved by a limited number of replicates.

Table of Contents

Introduction

Proper management and conservation of animal populations are highly dependent on reliable estimates of population size. A wide range of different methods for estimating population size exists, such as pellet group counts (Neff 1968) or counts along predetermined routes (Vincent *et al.* 1991). The most reliable method is regarded to be capture-mark recapture (CMR) techniques (Lebreton *et al.* 1992). The basic idea of CMR methods is that one catches a sample of animals from a population, marks them and releases them. Then after allowing the marked animals to become thoroughly mixed into the rest of the population, one takes another sample. The ratio of newly captured individuals in subsequent sessions can be used to estimate the size of the whole population. Various CMR approaches are available, each involving different assumptions or different statistical methods for arriving at an estimate of population size (Greenwood 1996). These census methods are often time consuming and expensive for research projects.

Over the last decade, advances in molecular techniques have made it possible to census animal populations by DNA analysis of non-invasive material such as faeces, urine, shed hair or feathers (e.g. Taberlet *et al.* 1997; Kohn *et al.* 1999, Ernest *et al.* 2000, Flagstad *et al.* 2004). In fact, sampling of such non-invasive material and subsequent identification of individuals by molecular tagging (Paalsbøll *et al.* 1997) is equivalent to the design of CMR. There are two possible outcomes of each sampling event. Either, the new sample may represent an individual that already has been encountered earlier or it represents a new individual. When adding more samples to the collection, the probability to come across a new individual will decrease and eventually reach zero. At this point the entire population has been sampled (Flagstad 2002). The use of a non-invasive approach may offer advantage over conventional CMR techniques, including increased capture probability, decreased tag loss, and potential to minimize the effects of capture and marking (Mills *et al.* 2000).

Dung and hair have been the most widely employed non-invasive material. Dung is of particular interest as all animals defecate regularly, and for many species, finding

dung is comparably simpler than finding the animal itself. Moreover, collection, storage, and transportation require little technology or expense. This source of DNA may be particularly useful in cases where the target species is vulnerable or endangered, and/or in cases where direct observation is difficult or time-consuming. Although the target animal may never be observed during studies based on non-invasive material, DNA analysis of such samples can provide answers to questions that earlier required direct observation and handling of animals. Indeed, there has been a steady accumulation of studies that have used molecular markers to identify individuals from non-invasively collected samples. In turn, this has allowed estimation of population size, inference of colonization history, detection of bottlenecks and assessment of other questions relevant to conservation and management of the target species (e.g. badger *Meles meles* (Frantz *et al.* 2003), brown bear *Ursus arctos* (Kohn *et al.* 1995; Taberlet *et al.* 1997; Wasser *et al.* 1997), chimpanzees *Pan troglodytes verus* (Morin *et al.* 2001), elephant *Loxodonta cyclotis* (Eggert *et al.* 2003), hartebeest *Alcelaphus buselaphus swaynei* (Flagstad *et al.* 2000), tiger *Panthera tigris amoyensis* (Wan *et al.* 2003) and wolverine *Gulo gulo* (Flagstad *et al.* 2004)

Non-invasive samples can give the same answers in genetic studies as analyses based on blood or tissue, but can be more problematic and time consuming due to the small amount of DNA or its poor quality (Gagneux *et al* 1997; Gerloff *et al.*1995; Taberlet and Luikart 1999). The main problem is the risk of false identification of individuals because of genotyping error. False identification can result from allelic dropout, where one of the two alleles of heterozygous individuals fails to amplify, thereby producing a false homozygote. False alleles, a less prominent genotyping error, are spurious amplification of alleles that do not exist (Taberlet *et al.* 1996). Allelic dropout have been reported in frequencies of >30% of the replicates that should have given a heterozygous profile (e.g. Gagneux *et al.*1997), and are doubtless the major source of genotyping error. To overcome the problem of large biases, which would occur with such false genotypes in the data set, an appropriate number of replicates is essential. Taberlet *et al.* (1996) proposed a method that would reduce the consequences of genotyping errors, the multiple-tube approach, where the principle is to replicate the genotyping experiments. The design of a suitable pilot project is also

recommended, as this will provide important information on the amount and nature of genotyping error (Taberlet *et al.* 1999). A pilot study may indicate whether reliable genotyping can be achieved with a reasonable number of replicates and whether the questions of interest can be answered through a purely non-invasive approach.

Norway has one of the largest continuous red deer (*Cervus elaphus*) populations in Europe (Langvatn 1999). The population increased considerably during the last century, as illustrated by hunting statistics. In 1894, only 84 animals were harvested, compared to about 25200 in 2003 (68% of licences issued, 96% in western Norway; Statistics Norway 2004). The current Norwegian population is estimated to count around 150,000 individuals, an estimate which simply is calculated as six times the number of animals harvested (R. Langvatn, pers. comm.).

Red deer inhabit forested landscapes often with very low visibility, making it difficult to assess population size by direct observation, except during the spring period when they gather on agricultural land. Therefore, conventional CMR-approaches have never been applied to estimate the population size in Norway. Rather, current census estimates are built on relatively scarce data collected from three different sources: (1) The “seen-deer” method, which is brief reports from hunters concerning hunting effort and geographic distribution of observed and shot animals (Veiberg *et al.* 2003). This method mainly aims at detecting whether population size is increasing or decreasing. (2) The “spring-counting” method, which detects animals that appear on farmed land in the spring (Veiberg *et al.* 2003). (3) Indirect data on individual performance (mainly body weight and ovulation rates) from hunted and road-killed animals (Langvatn *et al.* 2004; Mysterud *et al.* 2001). These indirect methods are based on changes in vital rates over time; decreasing performance suggests that population density is increasing. Census figures obtained from these data are obviously fairly rough, mainly assessing whether or not population size increase or decrease, providing a risk for over- or under-exploitation of harvested populations. Indeed, reliable estimates of population size are needed for appropriate management of this economically important species.

DNA analysis of feces may offer a good alternative to census red deer populations. However, since nobody has yet attempted to genotype DNA from red deer feces, it is necessary to select an appropriate set of genetic markers and optimize the methodology before any estimates of population size can be achieved by CMR methods. The primary aim for this study is to obtain genetic profiles from red deer feces, and evaluate to what extent different individuals can be reliably recognized from feces.

Material and methods

Samples

Parallel feces and blood samples were collected from individuals of red deer. A total of 42 complementary samples were included in the study. All the samples were collected from farmed animals at "The Norwegian Red deer centre", Kvalstad Gard on Svanøy in Sogn og Fjordane, Norway. The animals on this centre originate from wild Norwegian red deer.

All samples were collected from living animals during spring 2001. The faecal samples were collected either from rectum during blood collection or immediately after we had observed the particular individual defecating. In both cases the exact identification were known. The samples were picked up using disposal gloves and transferred into separate plastic bags. At this point the feces were completely fresh and warm. The plastic bags with the samples were immediately frozen at -20 °C, avoiding contamination from humans or other species' DNA. Before extraction the faecal samples were air dried at room temperature for 4-5 days, and subsequently stored dry at room temperature. Blood samples were stored at -20 °C until extraction.

DNA extraction

Feces

Extraction using Dynabeads

Variant 1

One dropping from each individual was put in a small sterile plastic-bag, then adding 400 µl phosphate-buffered saline (PBS) with pH 7.4. The surface of the feces was gently washed to release epithelial cells. The supernatant was transferred to a 1.5 mL eppendorf tube, and handled with 200 µl of Dynabeands DNA Direct (Dynal AS, Oslo, Norway) according to Rudi *et al.* (1997). This approach is identical to the method that was successfully used by Flagstad *et al*. (1999, 2000). When carrying out this procedure, I observed that most of the PBS was absorbed into the pellets, and that just a small amount of supernatant could be transferred to the next step of the

procedure. Given this observation, I decided to use also a larger amount of PSB, as described below in variant 2.

Variant 2

One dropping from each individual was put in a 50 ml falcon-tube, and approximately 4ml of PBS (pH = 7.4) was added. The tube was shaken at low speed for 6 min until the buffer had got a brownish colour. Importantly, the shaking was stopped before the feces dissolved. 800 µl of the supernatant was transferred to a 1,5 mL eppendorf tube, and handled with 200 µl of Dynabeads DNA Direct (Dynal AS, Oslo, Norway) according to Rudi *et al.* (1997).

Extraction using silica columns

Initial PCR amplifications indicated that the two variants of the Dynabeads approach, described above, had a rather low amplification success. Thus, I decided to test a Silica-based technique, which has been successfully used for carnivore faeces (Flagstad *et al.* 2004).

DNA from feces was isolated using the QIAamp DNA stool kit (GmbH, Hilden, Germany), following a slightly modified protocol from that provided by the manufacturer:

a) *Cutting feces and lysis of cells.* Small pieces (~200 mg) were cut from the outside of a dried feces-pellet and placed in a tube together with lysis buffer for about 50 minutes. After the sample had been homogenized and centrifuged the supernatant was transferred to a new tube.
b) *Removal of inhibitors.* To absorb compounds that could degrade and inhibit downstream enzymatic reactions in the extract, an InhibitEX tablet from the stool kit was added and allowed to incubate for one minute. Two repeated steps, including centrifugation and transferring of the supernatant into a new tube, were performed to remove excrement particles and the tablet matrix with bound inhibitors.
c) *Removal of proteins.* Proteins were removed by adding proteinase K followed by incubation at 70 °C for 10 minutes.

d) *Precipitation and elution of DNA.* After incubation, 96% ethanol was added to precipitate DNA The extract was loaded in portions onto a spin column and centrifuged. DNA-molecules were in this step adsorbed to the membrane of the spin column. After two washing steps DNA was finally eluted in 100 µl buffer.

Each round of extraction included 8-12 samples and a negative control. In order to reduce the risk of contamination, a separate laboratory was used during all DNA extraction from feces.

Blood

DNA from blood was extracted using a standard phenol:chloroform protocol (Sambrook *et al.* 1989)

Microsatellite genotyping

PCR optimization

PCR amplifications were performed in 10 µl volumes. In order to optimize amplification conditions for the different microsatellite markers (Table 1), I tested different quantities and concentration of PCR reagents as follows. The bovine serum albumin (BSA) was tested at 0 µg, 0.125 µg, 0.25 µg, 0.5 µg and 0.625 µg. Primers were tested at 2.56 pmol, 3.2 pmol, 4.096 pmol, 6.4 pmol and 12.8 pmol. $MgCl_2$ concentration was tested at 1.5 mM, 1.75 mM, 2.5 mM, 3.0 mM and 5.0 mM. Finally, the Hot Star Taq polymerase (Qiagen) was tested at 0.25units, 0.3units and 0.45units. Annealing temperatures between 50°C and 68°C were tested separately and by using gradients. However, satisfactory PCR conditions were first found by using a touchdown procedure; i.e. a PCR protocol where the annealing temperature is slightly reduced for each of the initial 10 or 15 cycles, followed by amplification for a number of cycles at the lowest temperature in the gradient. Three touchdown protocols were tested; 61°C - 48°C, 56°C - 50°C and 58°C - 52°C. In addition to these optimizations, different dilutions of the DNA extractions were tested; x1, x5, x10, and x15. After different combinations of these quantities, temperatures, and dilutions had been tested, optimal amplification conditions were found in a 10 µl reaction containing 3.0 mM

$MgCl_2$, 0.2 mM dNTP, 3.2 pmol of each primer, 0.5 μg of bovine serum albumin (BSA), 0.45 units of Hot Star Taq polymerase (Qiagen), and 2 μl of undiluted template. The PCR profile included an initial denaturation step of 95°C for 15 min, followed by 11 touch-down cycles with 94°C for 30 s, 58°C for 30 s decreasing 0.5°C each cycle and 72°C for 1 min, and 31 cycles of 94°C for 30 s, 52° C for 30 s and 72°C for 1 min. A final extension at 72°C for 10 min was added.

PCR amplification

As an initial test of the performance of the samples, amplification of nuclear DNA (nDNA) was carried out using one marker (BM4208; Bishop et al. 1994) that gives strong amplification products, clearly visible on an agarose gel. Amplifications were performed as described above, and visualized on a 2 % agarose gel. Samples that gave visible products on the gel were subsequently amplified with fluorescently labelled primers, using the optimal conditions and PCR profile as described above. All samples were run through several replicates for each locus, and a single-locus genotype was never accepted before it had shown two unambiguously identical heterozygous profiles or three identical homozygous profiles. Six microsatellite loci were used to achieve an appropriate resolution among different individuals (Table 1). All six markers were dinucleotide repeats with allele lengths ranging from 80-240 bp (Table 2).

Table 1 Primer sequences and references for the microsatellite markers used in this study.

Locus	Accession no.	Repeat sequence	Primer sequence (5'-3')	References
NVHRT 48	AF068214	$(GT)_2ATGTAT(GT)_6AT(GT)_{12}$	F: CGTGAATCTTAACCAGGTCT R: GGTCAGCTTCATTTAGAAAC	Roed & Midthjell (1998)
NVHRT 73	AF068218	$(CT)_4GT(CT)_3GCCTGT(CT)_2CCTT(CT)_3TT(CT)_{13}CACT(CA)_8TA(CA)_3$	F:CTTGCCCATTTAGTGTTTTCT R:TGCGTGTCATTGAATAGGAG	Roed & Midthjell (1998)
BM888	G18484	Not specified	F: AGGCCATATAGGAGGCAAGCTT R: CTCGGTGAGCTCAAAACGAG	Bishop *et al.* (1994)
BM4208	G18509	Not specified	F:TCAGTACACTGGCCACCATG R:CACTGCATGCTTTTCCAAAC	Bishop *et al.*(1994)
MCM 58	L34283	$(AC)_{25}$	F: CTGGGTCTGTATAAGCACGTCTCC R: CAGAACAATAAACGCTAAACCAGAGC	Hulme *et al.*(1994)
Oar FCB 304	LO1535	$(TC)_6GC(TC)_4GC(TC)_4GC(TC)_{11}(AC)_{15}$	F:CCCTAGGAGCTTTCAATAAAGAATCGG R:CGCTGCTGTCAACTGGGTCAGGG	Buchanan & Crawford (1993)

DNA from blood extracts was amplified using the same conditions and PCR profile as described above, but with a lower number of cycles (Table 2).

PCR products for both blood and faeces were in all cases run on a MegaBACE 1000 instrument (Molecular Dynamics, Amersham Bioscience 2001) and the subsequent determination of allele lengths was preformed with MegaBACE Genetic Profiler 1.5 (Amersham Bioscience 2001).

Data analysis

To determine the reliability of the results from the faeces samples, all multi-locus genotypes from feces were compared to those from blood from the same individuals. In this comparison, the assumption was made that the correct genotypes were those obtained from the blood samples.

Allelic dropout was interpreted in cases where at least one replicate showed a homozygous pattern while others were heterozygous. Alleles that occurred in only one of the independent replicates were interpreted as false alleles. A single-locus genotype was not accepted until it had shown three identical homozygous profiles or two identical heterozygous profiles.

Table 2 Basic parameters from PCR amplification of the microsatellite markers used in this study.

Locus	**Number of alleles**	**Range**	**Number of PCR-cycles, feces**	**Number of PCR-cycles, blood**	**PI_{rand}**	**PI_{sibs}**
NVHRT 48	5	84-107	11+31	11+25	0.153	0.451
NVHRT 73	4	210-234	11+31	11+25	0.391	0.649
BM888	9	207-237	11+26	11+25	0.078	0.383
BM4208	4	151-162	11+31	11+25	0.189	0.487
MCM 58	6	167-185	11+31	11+25	0.078	0.376
Oar FCB 304	8	130-149	11+31	11+28	0.094	0.399

The allele frequency data from the blood samples were calculated using the Excel Microsatellite Toolkit (an add-in utility for Microsoft® Excel, MS_toolkitt.xla; Park, 2001). The heterozygosity for each locus were calculated using the GENETIX 4.04 (Belkhir *et al.* 1999). The probability of identity (PI) for both unrelated individuals and full siblings was calculated based on the allele frequency (Table 3) using the program GEMINI 1.2.0 (Valiere *et al.* 2002). Several input files were made before using Gemini:

- Allele Frequency File (AF File)
- File containing the heterozygosity for each locus
- File containing observed error rates assuming variability in performance across loci but not across samples.
- File containing observed error rates assuming variability in performance across samples, but not across loci.
- File containing observed error rates assuming variability across loci and samples.
- Consensus file, containing all single-locus genotypes as inferred from at least two replicates.

Finally, Gemini was used to simulate the minimum number of replicates required to obtain a certain level of reliability for the consensus genotypes. This simulation determines the threshold, i.e. the minimum number of replicates, for which a single-locus consensus genotype can be accepted with a certain confidence.

Results

PCR amplifications from the Dynabeads extractions were unstable and inconsistent, and both variants of this approach showed extensive amplification failure as well as allelic dropout. Therefore, the Dynabeads approach was not used for the full set of samples and loci. Variant 1 was tested on three loci and a few samples. More than 60 % of the replicates failed to amplify, and successful amplifications showed an allelic dropout rate of approximately 30 % (Appendix I). Variant 2 was tested on 8 samples. However, the initial test indicated that only three of these could be amplified (Appendix II) and further testing was restricted to these samples. Amplification failed in 67 of the 72 replicates, and the correct genotype was only obtained in one of the five replicates that amplified successfully (Appendix III). Obviously, the extractions using Dynabeads did not gain DNA of sufficient quantity and/or quality for reliable genotyping.

In contrast, DNA of sufficient quality and/or quantity to be genotyped for all six loci was extracted from 23 of the 42 faces samples (54.8 %) by the silica-based method. All these samples were run through at least two replicates (between two and nine positive amplifications), and allelic dropout was found in 7.7 % of the independent replicates at heterozygous loci (Table 3). Table 3 shows that the presence of allelic dropout is concentrated around a few individuals. Only 8 samples showed dropout, four of which showed >20% dropout. False alleles were less prominent and occurred in only four individuals. Only 1.3 % of all amplifications showed false alleles (Table 3). A comparison of blood and feces samples from the same individual showed that correct consensus genotypes were obtained at all loci and for all 23 feces samples already after the three first replicates (Table 3; Figure 1).

As described above, a single-locus genotype was not accepted until it had shown three identical homozygous profiles or two identical heterozygous profiles. Table 4 shows amplification success and genotyping error for the 19 samples, which did not give DNA of sufficient quality and/or quantity to fulfill these criteria.

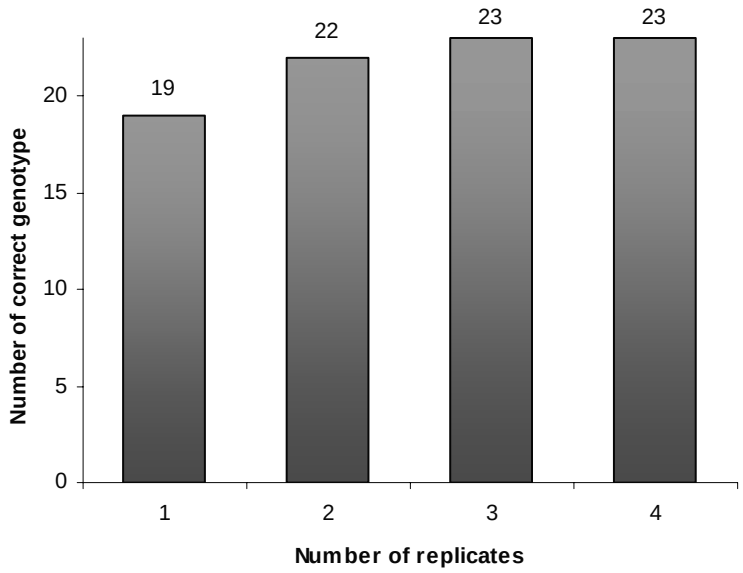

Figure 1 Cumulative number of red deer stool samples (amplifying samples only) showing the correct genotype at all loci after one, two, three and four replicates.

The main problem for these samples was amplification failure, but they also showed very high frequencies of allelic dropout and false alleles. One of the rather difficult samples to determine was number 19, for which all amplifications were correct at three loci. However, a large amount of false alleles and allelic dropout were observed in the remaining three loci (Table 4). Even with repeated extraction and genotyping of both faeces and blood from this individual, a reliable multi-locus genotype could not be obtained from this individual.

Table 3 Extraction using silica columns: Simplified genotypes from replicates of analysed faecal samples that **gave DNA of sufficient quality** to obtain reliable multi-locus genotypes. The outcome of each replicate is indicated by an integer corresponding to the number of amplified alleles (false alleles included). Replicates showing an incorrect genotype (as compared to that obtained from blood of the same individual) are indicated in bold and underlined. Bold 1 indicates an allelic dropout, 2* and 3* indicates a false allele, at a homozygote and a heterozygote locus, respectively.

Sample	BM888	OarFCB304	NVHRT 73	BM4208	MCM58	NVHRT 48	Average dropout+	Average false allele	Average based on loci
1	222	222	222	111	222	111	0.0 %	0.0 %	
2	111	111	222	222	222	222	0.0 %	0.0 %	
8	222	**1** 22	111	222	222	111	8.3 %	0.0 %	
10	222	22 **1**	222	222222	2 **11** 2 **1** 2	22222	15.4 %	0.0 %	
14	1111	222**11**2	2**1111112*1**	222**3***2**1**	11**2*2***11**2***	111111	47.6 %	13.2 %	
15	111	222	111	222	222	111	0.0 %	0.0 %	
20	111	1**2***11**2***1	111	111	111	111	0.0 %	5.7 %	
21	111	222	22	22	11	111	0.0 %	0.0 %	
22	222	222	222	111	111	22	0.0 %	0.0 %	
25	111	111	11111	111111	222222	222222	0.0 %	0.0 %	
26	111	222	222	222	222	111	0.0 %	0.0 %	
27	222	222	111	222	222	111	0.0 %	0.0 %	
30	**1**22	1111	11111	111	222	111	16.7 %	0.0 %	
31	222	222	111	222	222	222	0.0 %	0.0 %	
32	2**1**22**11**	1111	111	222	111	111	33.3 %	0.0 %	
33	1**2*2***111§	222	111	222	111	**11**22**1**2	25.0 %	8.3 %	
36	2222	2222	1111	2222	2222	2222	0.0 %	0.0 %	
38	111	222	1111	22222	22222	2222	0.0 %	0.0 %	
39	11111	111111	**1**2222	**2***1111	222222	1111	9.1 %	3.2 %	
40	22	22	11111	212	212	2212	21.4 %	0.0 %	
41	222	222	111	222	222	222	0.0 %	0.0 %	
43	22	22	111	222	222	222	0.0 %	0.0 %	
45	222	222	111	222	222	111	0.0 %	0.0 %	
Average dropout	9.8 %	7.5 %	25.8 %	3.4 %	6.3 %	9.3 %	7.7 %		8,9 %
Average false allele	2.6 %	3.8 %	1.2 %	2.4 %	3.6 %	0.0 %		1.3 %	2,2 %

+ The dropout rate was estimated only from heterozygosis loci in a particular individual

§ Repetition two and three in sample 33 showed two different heterozygote amplifications, thus more amplifications were needed to fulfil the predefined criteria

Table 4 Extraction using silica columns: Simplified genotypes from the replicates of analyzed faecal samples that **did not give DNA of sufficient** quality to obtain reliable multi-locus genotypes. Symbols and explanations are the same as in table 3. Replicates that failed to amplify are marked as #.

Sample	BM888	OarFCB304	NVHRT 73	BM4208	MCM58	NVHRT 48	Average dropout	Average false allele	Average failed amplified	Average based on loci
3	######	######	#**3***####	######	**2***##**2***#**2***	##**2***###	0,0 %	13,9 %	86,1 %	
4	### 1 # **2***	######	1 # 1111	# **1** ## **11**	##### **1**	# **3*** # **111**	38,9 %	5,6 %	58,3 %	
5	######	**3***#**3***###	######	#####**2***	#**3***#2##	**2***#####	0,0 %	13,9 %	83,3 %	
6	######	##**3***###	#**11**#**1**#	###**3***##	#**1**####	#**11**###	25,0 %	5,6 %	77,8 %	
7	###	###	###	#**1**#	###	###	5,6 %	0,0 %	94,4 %	
9	#**3***#	###	###	###	1##	###	0,0 %	5,6 %	88,9 %	
12	####**1**#	######	#####**3***	######	**2***11###	**1**##**1**##	12,5 %	5,6 %	80,6 %	
13	######	#**1**####	####**1**#	1#####	**13***####	#**3***#**11**#	20,8 %	5,6 %	77,8 %	
16	###	###	##**1**	###	#2**3***	###	8,3 %	5,6 %	83,3 %	
17	######	#### **2*** #	#### **2*** #	**2*** #####	**3*3*3*** # **3*3***	**2*** # **2*** ###	0,0 %	27,8 %	72,2 %	
18	##**3***###	###**1**#**1**	######	######	**2***#1#**2***#	######	8,3 %	8,3 %	80,6 %	
19	**2*2*2*** # **2*2***	222	111111	**111111113***	**#3*3*3* 1 3***	222	42,9 %	30,3 %	6,1 %	
24	11 ####	1 **2*2*** ###	1111 ##	**11**##**1**#	**3*3*3*11** 2	**1#1#1#**	44,4 %	13,9 %	41,7 %	
28	######	**1**2**1**###	1#1###	**1**22#**1**#	**3*1**#**13*3***	**1**#**13*3***#	26,7 %	13,9 %	50,0 %	
29	#2####	#**3***####	##2##**1**	11#**2***##	1#**2*2*2***#	######	5,6 %	13,9 %	69,4 %	
34	###	###	###	**2***##	**3*11**	###	13,3 %	11,1 %	77,8 %	
35	**1**#**1**	##**3***	**2***##	###	###	###	16,7 %	11,1 %	77,8 %	
37	**2***#**2***######	111111	# 1 # ### #**2***#	#11111	**1#13*3*3*#13***	**1111**2##2#	38,9 %	14,6 %	41,7 %	
44	#**3***####	**1**22#**11**	**1113*3*3***	**3*3*3*111**	**1**#2**113***	1##**2***#1	33,3 %	25,0 %	27,8 %	
Average dropout	6,3 %	14,0 %	15,0 %	39,2 %	20,8 %	26,1 %	18,0 %			20,2 %
Average false allele	10,8 %	8,3 %	8,3 %	9,4 %	34,4 %	9,1 %		12,2 %		13,4 %
Average failed amplified	83,3 %	69,8 %	68,8 %	65,6 %	45,8 %	65,7 %			67,1 %	66,5 %

Table 5 Allele frequencies for each locus in the red deer used in this thesis.

Locus	No of alleles	1	2	3	4	5	6	7	8	9
BM888	9	0.065	0.152	0.044	0.022	0.109	0.022	0.391	0.174	0.022
OarFCB304	8	0.217	0.413	0.087	0.109	0.109	0.022	0.022	0.022	
NVHRT 73	4	0.109	0.761	0.022	0.109					
BM4208	4	0.196	0.500	0.044	0.261					
MCM58	6	0.261	0.261	0.065	0.152	0.044	0.217			
NVHRT 48	5	0.044	0.261	0.130	0.478	0.087				

Probability of identity (PI)

A total of 36 alleles at the six applied loci were detected, with an average of 6 alleles per locus. The frequencies of all alleles are given in table 5. Calculated from the reliable genotypes of 23 red deer, the PI for unrelated individuals was 6.5×10^{-6}, whereas the PI was estimated to 8.0×10^{-3} for siblings. This shows that the 6 loci used in this study were sufficient to distinguish with 99% certainty between sibling red deer (Figure 2).

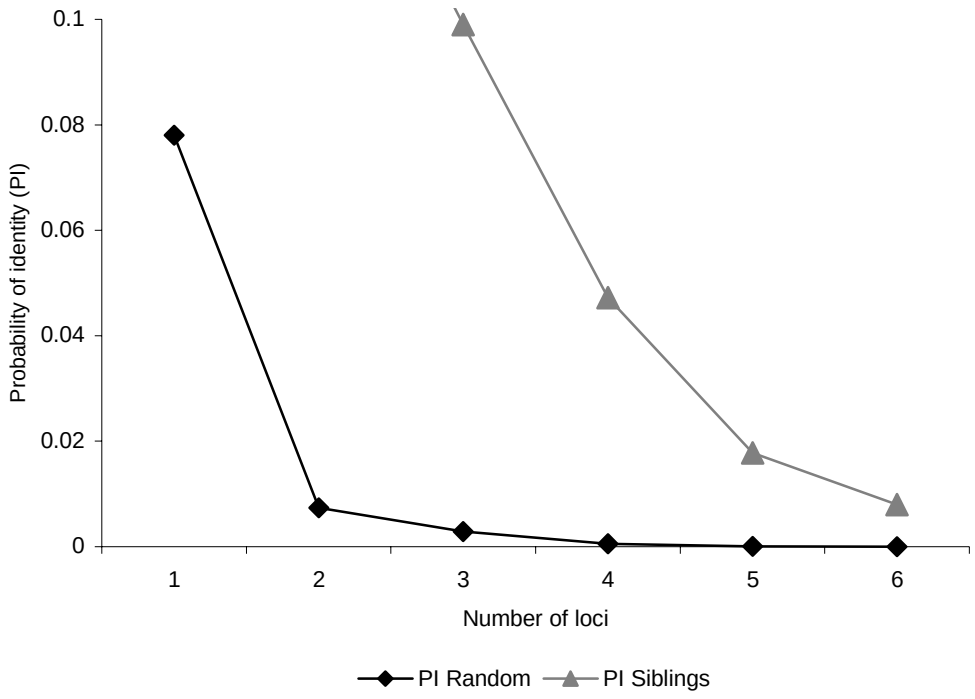

Figure 2 The probability of identity (PI) for samples taken from randomly chosen unrelated individuals and from siblings as a function of the number of loci used in the study.

Simulations

Using the observed data on error rates and variability in performance across loci and samples, one thousand simulations indicated that 85 % of the obtained multi-locus genotypes are expected to be correct already after one single replicate (Figure 3). As expected, this is in accordance with the empirical data, in which 19 of 23 samples (82.6 %) showed a correct multi-locus genotype after the first replicate. Most samples were run though at least three replicates (Table 3). According to the simulations, 95 % of the obtained multi-locus genotypes are expected to be correct after 3 replicates. In my case, all 23 samples showed a correct multi-locus genotype. However, given the observed variability in performance across samples and loci, the simulations show that an erroneous multi-locus genotype is expected to occur once for every 20 samples. Increasing the number of replicates to five, the expected error rate is reduced to 0.01.

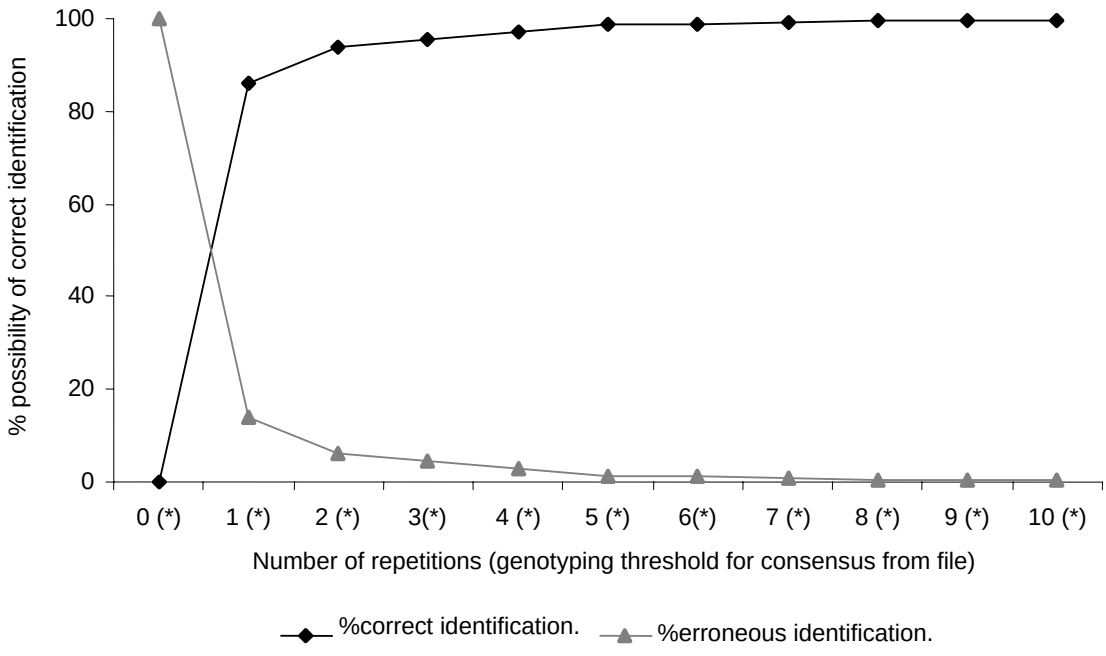

Figure 3 Average result of 1000 simulations of genotyping in Gemini, with threshold for consensus from a file constructed in the program.

Discussion

Amplification success - factors that may explain the poor performance of the Dynabeads methodology

Obtaining satisfactory DNA extracts is the main difficulty when using faecal sampling for genotyping. For example, PCR may fail because of degradation of DNA and/or the presence of inhibitors if the applied extraction protocol is inappropriate. In this study, the dynabeands method, which has been successfully used in other species (Flagstad *et al.* 1999; 2000), was shown to be inappropriate for red deer faeces. However, the stool kit method yielded extracts of good quality.

The amplification success of the latter method was 54.8 %, and compared to other studies this is an acceptable result. Earlier studies using faeces as the source of DNA have reported amplification success rates from 20 % (Dallas *et al.* 2003) to almost 100% (Flagstad *et al.* 1999). Studies where the faeces has been frozen have had success rates between 27% and 97.5% (e.g. (Frantzen *et al.* 1998; Hedmark *et al.* 2004; Huber *et al.* 2002; Lucchini *et al.* 2002).

The high variability in amplification success between the two different methodological approaches and different studies in general may be explained by a number of factors. For example, the way the samples are preserved could play a significant role. During sample storage it is vital that the opportunity for nucleases to degrade the DNA is minimized. This requires that the molecular environment of the DNA be physiologically inhospitable to enzymatic activity, which can be activated by either physical or chemical means. For feces, this means that the sample should be dehydrated; either by air drying (e.g. sheep *Ovis aries* and reindeer *Rangifer tarandus*; Flagstad *et al.* 1999), alcohol treatment (e.g. brown bear *Ursus arctos*; Murphy *et al.* 2002), frozen at -20 °C (e.g. mountain lion *Puma concolor*; Ernest *et al.* 2000), or saturated in a buffer containing high concentrations of salts or other chemicals that will inhibit damaging enzymes (e.g. DET buffer; Frantzen *et al.* 1998). Different studies have compared these and other preservation methods for stool samples and the overall outcome of the tests indicates that the best storage is in 70% ethanol (Frantz *et al.* 2003; Murphy *et al.* 2002).

A second factor that might affect amplification success is the diet of the target species. For example, faecal DNA genotyping from alpine ibex *Capra ibex* and Corsican mouflon *Ovis musimon* showed low error rates only when carried out on samples collected in winter (99% and 95% amplification of samples, respectively) as compared to spring, when only 52% and 59% of samples produced reliable genotypes. A possible explanation for this observation is seasonal dietary and/or climatic characteristics. Both ibex and mouflon eat young buds and shoots with minimal fibre during spring, which may decrease intestinal membrane abrasion and/or increase the passage of material, thus reducing the intestinal cell content in spring faeces (Piggott & Taylor 2003). Other seasonal effects have also been reported. For example, wolf samples collected fresh on snow in winter produced DNA extracts of higher quality as compared to those collected on bare ground during summer (Lucchini *et al.* 2002).

Although partly speculative, it is likely that the failure of the PBS-based surface-washing technique with subsequent DNA extraction by Dynabeads can be explained by a combination of the factors mentioned above.

First, the samples in this study may not have been stored optimally. We stored them frozen at -20°C and then later dried them. In comparison, Flagstad *et al*. (1999), who reported high amplification success, dried the droppings directly over night and stored them dry. Our initial freezing of the samples may have caused that the outer epithelial cells on the droppings have been washed off during the defrosting process. Thereby, the number of cells obtained using the PBS surface-washing technique could have been much smaller than in Flagstad *et al.* (1999).

Regarding diet, the red deer used in this project were farmed animals, eating bundled grass and concentrated cattle food in addition to grass, buds and shoots; i.e. relatively soft food. In contrast, the faeces samples from sheep and reindeer examined in Flagstad et al. (1999) were sampled in winter and had a diet consisting mainly of dry hay (Ø. Flagstad pers. comm.). It seems likely that the soft diet of the farmed red deer serves as a less efficient collector of epithelial cells than the dry and abrasive winter food of reindeer and sheep. Finally, although not experimentally tested, one can

imagine that the varied diet of red deer may contain a larger amount of PCR inhibitors than the simple winter food of reindeer and sheep. The Dynabeads approach does not include any step specially designed for removal of inhibitors, whereas the Silica-based method includes such a step. This may explain the relatively good results obtained by the latter approach as compared to the Dynabeads methodology.

Reliable genotyping from faeces

The rate of microsatellite genotyping error from non-invasive samples may depend on a range of biological and technical factors, as described above in relation to amplification success. The allelic dropout rate observed herein was 7.7%, which was slightly higher than that of 2 % reported in Flagstad *et al.* (1999). However, the error rate was not high as compared to other studies involving crushing or homogenizing whole or partial faeces (e.g. mountain lions *Puma concolor*, 8%, Ernest *et al.*2000; wolves *Canis lupus* 18%, Lucchini *et al.*2002).

My pre-defined scoring criteria were three unambiguously identical homozygous profiles or two identical heterozygous profiles before acceptance of a single-locus consensus genotype. When simulating the minimum number of typing repetitions required to obtain identification reliability at a certain level, 1000 simulations indicated that 4-5 repetitions were required to obtain a 98-99% probability of correct identification (Figure 3). However, my comparison of 23 blood and faeces samples demonstrated that correct identification was already obtained after the three first replicates for all samples (Table 4; Figure 1). Nevertheless, given the observed variability in performance, in particular across samples, the simulations showed that 3 replicates per locus and sample would give only 95% correct identification in the long run – i.e. that an erroneous multi-locus genotype would occur once for every 20 samples. This result suggests that samples of different quality should be treated with different criteria for accepting single-locus genotypes. Three replicates should be enough for samples showing no or little error, as is the case for most of the 23 samples in this study. However, samples that amplify readily, but show relatively high amounts of genotyping error (e.g. sample 14 and 33, Table 4), should probably be replicated in at least five replicates. Such a strategy, as recommended by Miller et

al. (2002), would contribute to minimise the occurrence of false genotypes in the data set.

Indeed, to avoid all the pit-falls associated with non-invasive genotyping from low copy number sources of DNA, it is important to carry out a pilot study prior to initiation of large-scale genotyping. A pilot project, as the one described herein, will contribute to quantification of error rates associated with different regimes of field sampling, as well as storage, extraction, and amplification protocols (Taberlet & Luikart 1999). Moreover, a pilot study will provide the necessary data for cost-benefit analyses to determine whether a non-invasive approach is feasible. Importantly, development of computer programs over the latest years has made it possible to simulate and quantify the effects of particular error rates on the outcomes of non-invasive genotyping, the associated analyses, and finally their subsequent biological interpretation (e.g. Gemini; Valiere *et al.*2002)

Probability of identity (PI)

Given the allele frequencies obtained in this study, the probability that two unrelated individuals share a multi-locus genotype is 6.5×10^{-6}. The corresponding figure for siblings is 8.0×10^{-3}. Taberlet and Luikart (1999) emphasized that the observed PI will fall somewhere between the two theoretical PIs (PI_{rand} and PI_{sibs}). The theoretical equations for PI will underestimate the true probability of finding identical genotypes in many natural populations (Taberlet & Luikart 1999). In contrast, PI for siblings represents the upper limit for PI, and may serve as a conservative guideline for the number of loci needed to have a high probability of resolving all individuals. Determining the PI that is 'sufficiently low' depends on the abundance of siblings in the population and the severity of the consequences of not being able to differentiate among all individuals. For capture-recapture studies a sufficiently low PI may be only 0.01 (Taberlet & Luikart 1999). In our study, we calculated the PI for siblings to be 0.008. Thus, given that levels of genetic variability is approximately the same in natural populations of red deer (Skog 2003) as in the farmed populations examined herein, six loci should be sufficient to distinguish between individuals also in natural populations. These data could subsequently be used in a CMR approach to estimating population size.

What can molecular genetic analyses add to traditional census methods?

Traditional methods of estimating animal abundance are based on direct counts of individuals (either free-ranging or captured animals), or indirect signs such as footprints and faeces. While direct approaches are effective for many animals, they are inadequate for species that are elusive and/or difficult to trap, and for endangered species for which such methods may be disturbing or even harmful.

There are also a variety of analytical disadvantages to abundance estimation based on trapping. One is that most trapping techniques are unable to provide 'snap-shot' estimates of population size for many species because they require many months or even years to obtain sufficient sample sizes (Kohn *et al.* 1999). In contrast, a relatively short time of intense sampling of faeces across a defined geographic area may give a sufficient sample size to estimate population size in that area (Eggert et al. 2003, Flagstad et al. 2004). Poor trapping success in conventional CMR-methods may also lead to underestimation of population size. For example, in a non-invasive approach to estimating population size in coyotes, Kohn *et al.* (1999) showed that more than two-third of the current population may have been missed by long-term ecological surveys, perhaps in part due to a low overall trapping efficiency of only one animal per 58 trap-nights.

Conventional faecal analyses can provide data on diet composition or parasite infestation, and novel methods of faecal hormone metabolite measurement allow assessment of reproductive state and physiological stress (Kohn, Wayne 1997). As recently shown by Huber *et al.* (2003b) faeces can also be used for measuring concentration of cortisol metabolites of red deer. By combining these methods with genotyping, it is possible to identify the individuals and determine their stress level related to environmental changes (Huber *et al.* 2003a; 2003b; 2003c). Another possible application of faecal DNA technology is identification of plants and animals consumed by the target species. In cases where food is thoroughly digested or difficult to identify, molecular identification will be invaluable.

Techniques are not yet developed to the stage where they reliably and cost-effectively can replace capture and sampling of animals. Some applications may not justify the expense and effort required. In this perspective, it is important to perform pilot studies in the laboratory and the field, and to do cost-benefit analyses before starting large-scale sampling. However, the cost for non-invasive genetic sampling is decreasing as the techniques are improved. Thus, such techniques will be an excellent alternative for species where current methods are not optimal for efficiently obtaining results or for animals where capture is impracticable. Three recent studies can be mentioned in this connection: a study of elephants (Eggert *et al.* 2003), one on wolverines (Flagstad *et al.* 2004) and finally a study of wolf packs in the Italian Alps (Lucchini *et al.* 2002). These three studies were solely based on molecular analyses of faeces, and contributed with new important information on population size and genetic structure.

As mentioned earlier, the indirect methods used today for estimating population size in Norwegian red deer are mainly aimed at determining whether or not population density increase or decrease, making it difficult to know how much can be harvested to maintain the population at a sustainable level. The limitations of current methods can be overcome by using genotyping from faeces to estimate the correct population size, thereby assessing the likely inaccuracy of the currently applied census methods. Such an approach would include sampling in a few areas of limited geographical size, representative for different parts of the distribution range in Norway. These data could in turn be used for reliable estimate of population size in the sampled areas, and thus provide a basis for inference of the total population size.

Conclusion

This study of red deer shows that DNA extracted from faeces is a viable source for determining individual genotypes, and that it can be obtained by a limited number of replicates. Thus, such non-invasive genetic sampling could contribute significantly to population dynamic studies by providing reliable estimates of crucial parameters like population size and immigration rate. Although we established a technique for obtaining red deer genetic profiles from faeces, further work is needed to design more efficient and appropriate sampling methods, and to design strategies to obtain census estimates across various spatial or temporal scales and population densities.

References

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1999) GENETIX 4.04, logiciel sous Windows TM pour la génétique des populations. Laberatoire Génôme et populations, CNRS UPR 9060, Montpellier, France.

Bishop MD, Kappes SM, Keele JW, *et al.* (1994) A Genetic-Linkage Map for Cattle. *Genetics* **136**, 619-639.

Buchanan FC, Crawford AM (1993) Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. *Anim Genet* **24**, 145.

Dallas JF, Coxon KE, Sykes T, *et al.* (2003) Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter Lutra lutra. *Molecular Ecology* **12**, 275-282.

Eggert LS, Eggert JA, Woodruff DS (2003) Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. *Molecular Ecology* **12**, 1389-1402.

Ernest HB, Penedo MCT, May BP, Syvanen M, Boyce WM (2000) Molecular tracking of mountain lions in the Yosemite Valley region in California: genetic analysis using microsatellites and faecal DNA. *Molecular Ecology* **9**, 433-441.

Flagstad Ø (2002) *Population history and non-invasive monitoring: Use of low copy number DNA in conservation genetics* Ph.D. thesis, University of Uppsala.

Flagstad Ø, Hedmark E, Landa A, *et al.* (2004) Colonization history and noninvasive monitoring of a re-established wolverine population. *Conserv. Biol.* **18**, 676-688.

Flagstad Ø, Roed K, Stacy JE, Jakobsen KS (1999) Reliable noninvasive genotyping based on excremental PCR of nuclear DNA purified with a magnetic bead protocol. *Molecular Ecology* **8**, 879-883.

Flagstad Ø, Syvertsen PO, Stenseth NC, *et al.* (2000) Genetic variability in Swayne's hartebeest, an endangered antelope of Ethiopia. *Conservation Biology* **14**, 254-264.

Frantz AC, Pope LC, Carpenter PJ, *et al.* (2003) Reliable microsatellite genotyping of the Eurasian badger (Meles meles) using faecal DNA. *Molecular Ecology* **12**, 1649-1661.

Frantzen MAJ, Silk JB, Ferguson JWH, Wayne RK, Kohn MH (1998) Empirical evaluation of preservation methods for faecal DNA. *Molecular Ecology* **7**, 1423-1428.

Gagneux P, Boesch C, Woodruff DS (1997) Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology* **6**, 861-868.

Gerloff U, Schlotterer C, Rassmann K, *et al.* (1995) Amplification of Hypervariable Simple Sequence Repeats (Microsatellites) from Excremental DNA of Wild Living Bonobos (Pan-Paniscus). *Molecular Ecology* **4**, 515-518.

Greenwood JJD (1996) Basic techniques. In: *Ecological census techniques : a handbook* (ed. Sutherland WJ). Cambridge University Press, Cambridge.

Hedmark E, Flagstad Ø, Segerström P, *et al.* (2004) DNA-based individuals and sex identification from wolverine (*Gulo gulo*) faeces and urine. *Conservation Genetics*, (in press).

Huber S, Bruns T, Arnold R (2003a) Genotyping herbivore feces facilitating their further analyses. *Wildlife Society Bulletin* **31**, 692-697.

Huber S, Bruns U, Arnold W (2002) Sex determination of red deer using polymerase chain reaction of DNA from feces. *Wildlife Society Bulletin* **30**, 208-212.

Huber S, Palme R, Arnold W (2003b) Effects of season, sex, and sample collection on concentrations of fecal cortisol metabolites in red deer (Cervus elaphus). *General and Comparative Endocrinology* **130**, 48-54.

Huber S, Palme R, Zenker W, Mostl E (2003c) Non-invasive monitoring of the adrenocortical response in red deer. *Journal of Wildlife Management* **67**, 258-266.

Hulme DJ, Silk JP, Redwin JM, Barendse W, Beh KJ (1994) Ten polymorphic ovine microsatellites. *Anim Genet* **25**, 434-435.

Kohn M, Knauer F, Stoffella A, Schroder W, Paabo S (1995) Conservation Genetics of the European Brown Bear - a Study Using Excremental Pcr of Nuclear and Mitochondrial Sequences. *Molecular Ecology* **4**, 95-103.

Kohn MH, Wayne RK (1997) Facts from feces revisited. *Trends in Ecology & Evolution* **12**, 223-227.

Kohn MH, York EC, Kamradt DA, *et al.* (1999) Estimating population size by genotyping faeces. *Proceedings of the Royal Society of London Series B-Biological Sciences* **266**, 657-663.

Langvatn R (1999) Hjortens erobring av Norge. In: *Brennpunkt Natur 98/99* (ed. KH B), pp. 49-70. Tapir, Trondheim, Norway.

Langvatn R, Mysterud A, Stenseth NC, Yoccoz NG (2004) Timing and synchrony of ovulation in red deer constrained by short northern summers. *American Naturalist* **163:**, 763-772.

Lebreton JD, Burnham KP, Clobert J, Anderson DR (1992) Modeling Survival and Testing Biological Hypotheses Using Marked Animals - a Unified Approach with Case-Studies. *Ecological Monographs* **62**, 67-118.

Lucchini V, Fabbri E, Marucco F, *et al.* (2002) Noninvasive molecular tracking of colonizing wolf (Canis lupus) packs in the western Italian Alps. *Molecular Ecology* **11**, 857-868.

Miller CR, Joyce P, Waits LP (2002) Assessing allelic dropout and genotype reliability using maximum likelihood. *Genetics* **160**, 357-366.

Mills LS, Citta JJ, Lair KP, Schwarz MK, Tallmon DA (2000) Estimating animal abundance using noninvasive DNA sampling: Promise and pitfalls. *Ecological Applications* **10**, 283-294.

Morin PA, Chambers KE, Boesch C, Vigilant L (2001) Quantitative polymerase chain reaction analysis of DNA from noninvasive samples for accurate microsatellite genotyping of wild chimpanzees (Pan troglodytes verus). *Molecular Ecology* **10**, 1835-1844.

Murphy MA, Waits LP, Kendall KC, *et al.* (2002) An evaluation of long-term preservation methods for brown bear (Ursus arctos) faecal DNA samples. *Conservation Genetics* **3**, 435-440.

Mysterud A, Yoccoz NG, Stenseth NC, Langvatn R (2001) Effects of age, sex and density on body weight of Norwegian red deer: evidence of density-dependent senescence. *Proceedings of the Royal Society of London Series B-Biological Sciences* **268**, 911-919.

Neff DJ (1968) Pellet-Group Count Technique for Big Game Trend Census and Distribution - a Review. *Journal of Wildlife Management* **32**, 597-614.

Park SDE (2001) The Excel Microsatellite Toolkit. In: *Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection [Ph.D. thesis]*, pp. An add-in utility for Microsoft® Excel (PC versions) that contains tools for population geneticists working with microsatellites. University of Dublin, Dublin.

Piggott MP, Taylor AC (2003) Remote collection of animal DNA and its applications in conservation management and understanding the population biology of rare and cryptic species. *Wildlife Research* **30**, 1-13.

Roed KH, Midthjell L (1998) Microsatellites in reindeer, Rangifer tarandus, and their use in other cervids. *Molecular Ecology* **7**, 1773-1776.

Rudi K, Kroken M, Dahlberg OJ, *et al.* (1997) Rapid, universal method to isolate PCR-ready DNA using magnetic beads. *Biotechniques* **22**, 506-511.

Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Skog A (2003) *A phylogeographic study of red deer (Cervus elephus) in Europe*. Cand. scient. thesis, University of Oslo.

StatisticsNorway (2004) Official hunting statistics of Norway 2003. Red deer hunting, by county. Statistics Norway, Oslo - Kongsvinger [also available at URL:http://www.ssb.no/english/subjects/10/04/10/hjortejakt_en/].

Taberlet P, Camarra JJ, Griffin S, *et al.* (1997) Noninvasive genetic tracking of the endangered Pyrenean brown bear population. *Molecular Ecology* **6**, 869-876.

Taberlet P, Griffin S, Goossens B, *et al.* (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189-3194.

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. *Biological Journal of the Linnean Society* **68**, 41-55.

Taberlet P, Waits LP, Luikart G (1999) Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution* **14**, 323-327.

Valiere N, Berthier P, Mouchiroud D, Pontier D (2002) GEMINI: software for testing the effects of genotyping errors and multitubes approach for individual identification. *Molecular Ecology Notes* **2**, 83-86.

Veiberg V, Meisingset EL, Samdal B (2003) Evaluering av Sett-hjort og vårteljing som hjelpemiddel for den lokale bestandsforvaltninga av hjort. Norsk Hjortesenter Fagrapport 1/03: 1-48.

Vincent JP, Gaillard JM, Bideau E (1991) Kilometric Index as Biological Indicator for Monitoring Forest Roe Deer Populations. *Acta Theriologica* **36**, 315-328.

Wan QH, Fang SG, Chen GF, *et al.* (2003) Use of oligonucleotide fingerprinting and faecal DNA in identifying the distribution of the Chinese tiger (Panthera tigris amoyensis Hilzheimer). *Biodiversity and Conservation* **12**, 1641-1648.

Wasser SK, Houston CS, Koehler GM, Cadd GG, Fain SR (1997) Techniques for application of faecal DNA methods to field studies of Ursids. *Molecular Ecology* **6**, 1091-1097.

Appendix

Appendix I Dynabeads variant 1: Simplified genotypes from some of the replicates of analyzed faecal samples extracted using variant 1. Only three loci were tested. The outcome of each replicate is indicated by an integer corresponding to the number of amplified alleles (false alleles included). Replicates showing an incorrect genotype (as compared to that obtained from blood of the same individual) are indicated in bold and underlined. Bold **1** indicates an allelic dropout, **2*** and **3*** indicates a false allele at a homozygous and a heterozygous locus, respectively. Replicates that failed to amplify are marked as #.

Sample	BM888	BM4208	MCM58	Average dropout	Average false allele	Average failed amplified	Average correct amplified	Average based on loci
1	###	###	**113***	33,3 %	11,1 %	66,7 %	0,0 %	
4	#**2*** 1	# **11**	**111**	83,3 %	11,1 %	22,2 %	0,0 %	
8	######	##22##	######	0,0 %	0,0 %	88,9 %	11,1 %	
10	######	##### 1	##### **3***	0,0 %	5,6 %	88,9 %	5,6 %	
11	###	###	###	0,0 %	0,0 %	100,0 %	0,0 %	
14	###	###	**2*** 1 #	0,0 %	11,1 %	77,8 %	11,1 %	
17	## 1	###	**# 11**	66,7 %	0,0 %	66,7 %	11,1 %	
19	**###**	# **11**	# **11**	66,7 %	0,0 %	55,6 %	0,0 %	
22	###	###	###	0,0 %	0,0 %	100,0 %	0,0 %	
25	###	###	###	0,0 %	0,0 %	100,0 %	0,0 %	
27	**11** 2	2 **11**	**111**	77,8 %	0,0 %	0,0 %	22,2 %	
28	###	###	**3*** ##	0,0 %	11,1 %	88,9 %	0,0 %	
30	######	######	#### **3*3***	0,0 %	11,1 %	88,9 %	0,0 %	
31	**1** ##	###	###	16,7 %	0,0 %	88,9 %	0,0 %	
33	111	**111**	## 1	100,0 %	0,0 %	22,2 %	44,4 %	
36	**3* 1 1**	**111**	222	55,6 %	11,1 %	0,0 %	33,3 %	
41	# **11**	## **1**	# **3* 3***	33,3 %	33,3 %	44,4 %	0,0 %	
Average dropout	16,7 %	43,3 %	23,5 %	31,4 %				27,8 %
Average false allele	3,3 %	0,0 %	13,3 %		6,2 %			5,6 %
Average failed amplified	75,0 %	66,7 %	58,3 %			64,7 %		66,7 %
Average correct amplified	10,0 %	6,7 %	8,3 %				8,2 %	8,3 %

Appendix II Dynabeads variant 2: Initial test of samples extracted using variant 2 were amplification products were visualized on a 2 % agarose gel. + indicates that amplification were visual on gel, and – indicates no visual amplification.

	Dilution of DNA extracts		
Sample	1x	5x	10x
3	+	-	+
6	-	-	-
7	+	-	-
8	-	-	+
15	-	-	-
28	-	-	-
35	-	-	-
36	-	-	+

Appendix III Dynabeads variant 2: Simplified genotypes from some of the replicates of analyzed faecal samples extracted using variant 2. Symbols and explanations are the same as in appendix I.

Sample	BM888	OarFCB304	NVHRT 73	BM4208	MCM58	NVHRT 48	Average dropout	Average false allele	Average failed amplified	Average correct amplified	Average based on loci
3	###	###	###	###	###	###	0,0 %	0,0 %	100,0 %	0,0 %	
7	###	**3*** ##	###	###	###	###	0,0 %	4,2 %	95,8 %	0,0 %	
8	###	###	###	###	# **3*3***	#1#	0,0 %	8,3 %	87,5 %	4,2 %	
36	###	###	###	###	# **3*** #	###	0,0 %	4,2 %	95,8 %	0,0 %	
Average dropout	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %				0,0 %
Average false allele	0,0 %	8,3 %	0,0 %	0,0 %	25,0 %	0,0 %		4,2 %			5,6 %
Average failed amplified	100,0 %	91,7 %	100,0 %	100,0 %	75,0 %	91,7 %			94,8 %		93,1 %
Average correct amplified	0,0 %	0,0 %	0,0 %	0,0 %	0,0 %	8,3 %				1,0 %	1,4 %